

Titre du projet : Discovery of new ARSACS modifier genes in the Saguenay founder population
Chercheur : Simon Girard

Rapport du projet

L'ataxie spastique autosomique récessive de Charlevoix-Saguenay (ARSACS) est prévalente dans la région du Saguenay–Lac-Saint-Jean (SLSJ) au Québec, en raison d'un fort effet fondateur et d'une expansion rapide de la population. L'une des variantes causales du gène *SACS*, c.8844delT, est retrouvée chez 94 % des patients du SLSJ, et est probablement la seule dont la fréquence a augmenté en raison de l'effet fondateur du SLSJ.

Les objectifs spécifiques de cette subvention étaient :

1. Génotyper les patients atteints d'ARSACS,
2. Obtenir des données généalogiques,
3. Combiner ces données avec des données phénotypiques afin d'identifier des gènes modificateurs.

Les données généalogiques ont été obtenues pour 166 patients (objectif 2), et le génotypage a été réalisé pour 71 de ces patients (objectif 1). Les variants spécifiques associés à l'ARSACS chez les patients étaient également disponibles grâce à des analyses de laboratoire clinique précédentes. L'analyse généalogique a été réalisée à l'aide du package R Genlib. Les segments identiques par descendance (IBD) ont été inférés à partir des génotypes en utilisant refinedIBD. Les haplotypes ont également été définis à partir des génotypes, autour du variant causal et sur l'ensemble du génome, en utilisant le package geneHapR.

En utilisant les haplotypes de différentes longueurs autour du variant causal et les informations cliniques sur les variants, nous avons pu identifier l'haplotype fondateur et le retracer dans les généalogies jusqu'au couple ayant la plus grande contribution aux porteurs de l'ARSACS (parents des patients porteurs du variant c.8844delT). Avec ce couple identifié, il sera possible, par le biais de simulations dans la généalogie, d'estimer un taux de porteurs attendu et une prévalence de la maladie dans les différents villages du SLSJ. Cette approche pourrait permettre de cibler certaines régions pour des tests de dépistage des porteurs.

La proportion de partage IBD entre toutes les paires de patients présente un pic sur le chromosome 13, correspondant au variant c.8844delT, connu pour causer l'ARSACS, comme attendu pour une maladie fondatrice. Aucun autre pic d'IBD n'a été observé ailleurs sur le génome dans le cadre d'une définition large du phénotype ARSACS qui aurait pu révéler des gènes modificateurs partagés (objectif 3).

En lien avec le partage IBD, des haplotypes de différentes longueurs ont été définis autour du variant ARSACS, et la phylogénie des haplotypes a été représentée pour cette

région génomique, reflétant ainsi les voies généalogiques par lesquelles le variant a évolué. L'haplotype le plus prévalent, associé au variant c.8844delT, était partagé par 86 % des patients, ce qui est cohérent avec les données cliniques rapportées précédemment. Nous sommes donc en mesure de conclure que plus de 85% des patients ARSACS porteur de la mutation c.8844delT proviennent d'un couple fondateur qui est maintenant connu. Cet haplotype pourrait être utilisé dans de futurs protocoles cliniques pour dépister les porteurs de l'ARSACS dans la population générale à faible coût.

Bien que la subvention soit terminée, les analyses des données sont toujours en cours au sein de notre équipe de recherche. Nous prévoyons utiliser les données haplotypiques, ainsi que les relations familiales incluses dans les génotypes et un groupe témoin issu de la population générale (base de données CARTaGENE), afin d'identifier des régions génomiques susceptibles de révéler des gènes modificateurs de l'ARSACS chez les patients. Pour ce faire, nous utiliserons le logiciel GCTA afin de réaliser une analyse d'association génomique basée sur les haplotypes, en intégrant les relations familiales comme effet aléatoire. L'utilisation des haplotypes permet d'éliminer le besoin de séquençage complet du génome, qui est plus coûteux. Nous commencerons avec une définition binaire large du phénotype ARSACS, puis nous explorerons plus en profondeur les sous-groupes de patients présentant des phénotypes plus spécifiques, dans le but d'identifier des gènes modificateurs pouvant influencer certaines caractéristiques de la maladie (objectif 3).

Le soutien de la fondation ARSACS a été crucial pour initier ce nouveau projet au sein du laboratoire Genopop. La modélisation complète de la transmission de l'ARSACS au sein de la population du SLSJ nous permet de disposer d'un modèle de suivi temporel et spatial d'une maladie rare au sein d'une population. Ce modèle est l'un des plus complet répertorié jusqu'à maintenant. **Ces travaux de recherche, combiné à ceux sur d'autres maladies rares, ont été des leviers importants qui ont permis l'octroi d'une importante subvention Génome Québec (annonce lors de la journée des maladies rares 2025).**

Résumé vulgarisé

L'ARSACS est une maladie génétique rare très présente au Saguenay–Lac-Saint-Jean en raison d'un effet fondateur : une mutation spécifique s'est propagée dans la population au fil des générations. Ce projet vise à reconstituer l'histoire génétique de cette mutation, identifier d'éventuels gènes modificateurs influençant la maladie, et améliorer le dépistage des porteurs.

Principaux résultats

- 166 patients ont été étudiés sur le plan généalogique et 71 ont été génotypés.
- Le couple fondateur ayant transmis la mutation pour la plupart des patients a été identifié, ce qui permettrait d'estimer la fréquence des porteurs dans la région.

- Un haplotype fondateur a été trouvé chez 86 % des patients, ce qui nous permettra d'améliorer les méthodes de dépistage déjà en place.

Ces découvertes aideront à mieux comprendre l'ARSACS, à optimiser le dépistage et à identifier de nouveaux facteurs génétiques pouvant influencer la maladie.

Project Title: Discovery of new ARSACS modifier genes in the Saguenay founder population

PI: Simon Girard

Project Report

Autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS) is highly prevalent in the Saguenay–Lac-Saint-Jean (SLSJ) region of Quebec due to a strong founder effect and rapid population expansion. One of the causal variants in the *SACS* gene, c.8844delT, is found in 94% of SLSJ patients and is likely the only one whose frequency has increased due to the SLSJ founder effect.

The specific objectives of this grant were to:

1. Genotype patients with ARSACS,
2. Obtain genealogical data,
3. Combine these data with phenotypic information to identify modifier genes.

Genealogical data were obtained for 166 patients (Objective 2), and genotyping was performed for 71 of these patients (Objective 1). The specific variants associated with ARSACS in the patients were also available through previous clinical laboratory analyses. Genealogical analysis was conducted using the *Genlib* R package. Identical-by-descent (IBD) segments were inferred from genotypes using *refinedIBD*. Haplotypes were also defined from the genotypes, both around the causal variant and across the entire genome, using the *geneHapR* package.

By analyzing haplotypes of different lengths around the causal variant and incorporating clinical variant data, we identified the founder haplotype and traced it back through genealogies to the couple who contributed the most to ARSACS carriers (parents of patients carrying the c.8844delT variant). With this identified couple, it will be possible to estimate the expected carrier rate and disease prevalence in different villages of the SLSJ through simulations within the genealogy. This approach could help target specific regions for carrier screening programs.

The proportion of IBD sharing between all patient pairs shows a peak on chromosome 13, corresponding to the c.8844delT variant, which is known to cause ARSACS—consistent with what is expected for a founder disease. No other IBD peaks were observed elsewhere in the genome when applying a broad definition of the ARSACS phenotype that could have revealed shared modifier genes (Objective 3).

Regarding IBD sharing, haplotypes of different lengths were defined around the ARSACS variant, and a haplotype phylogeny was constructed for this genomic region, reflecting the genealogical pathways through which the variant has evolved. The most prevalent haplotype, associated with the c.8844delT variant, was shared by 86% of patients, aligning with previously reported clinical data. We can thus conclude that more than 85% of ARSACS patients carrying the c.8844delT mutation originate from a known founder couple. This haplotype could be used in future clinical protocols to screen for ARSACS carriers in the general population at a lower cost.

Although the grant period has ended, data analyses are still ongoing within our research team. We plan to use haplotype data, as well as familial relationships embedded in genotypes and a control group from the general population (CARTaGENE database), to identify genomic regions that may reveal ARSACS modifier genes in patients. To achieve this, we will use the *GCTA* software to conduct a haplotype-based genome-wide association study (GWAS), incorporating familial relationships as a random effect. Using haplotypes eliminates the need for whole-genome sequencing, which is more expensive. We will begin with a broad binary definition of the ARSACS phenotype and then explore subgroups of patients with more specific phenotypes to identify modifier genes that may influence particular disease characteristics (Objective 3).

Support from the ARSACS Foundation has been crucial in initiating this new project within the *Genopop* laboratory. The comprehensive modeling of ARSACS transmission within the SLSJ population provides us with a temporal and spatial monitoring model for a rare disease within a population. This model is among the most complete documented to date. This research, combined with studies on other rare diseases, has been a significant catalyst in securing a major grant from Génome Québec (announcement scheduled for Rare Disease Day 2025).

Lay Summary

ARSACS is a rare genetic disease highly prevalent in the Saguenay–Lac-Saint-Jean region due to a founder effect—a specific mutation that has spread through the population over generations. This project aims to reconstruct the genetic history of this mutation, identify potential modifier genes influencing the disease, and improve carrier screening.

Key Findings

- **166 patients** were studied genealogically, and **71** were genotyped.

- The **founder couple** responsible for transmitting the mutation in most patients has been identified, allowing us to estimate the carrier frequency in the region.
- A **founder haplotype** was found in **86%** of patients, which will help improve existing screening methods.

These discoveries will enhance our understanding of ARSACS, optimize screening efforts, and identify new genetic factors that may influence disease progression.